



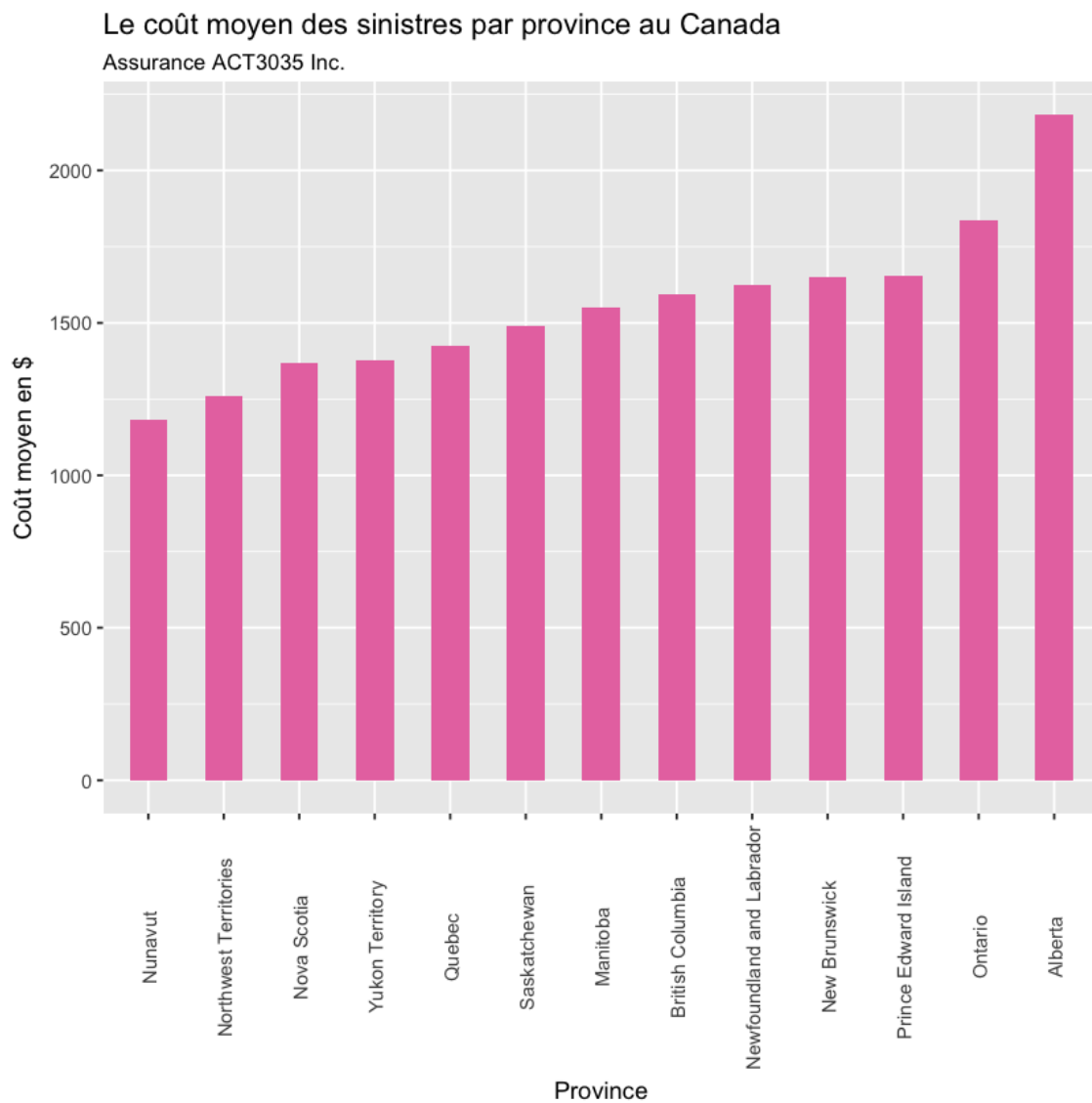
Question 6 (6 pts)

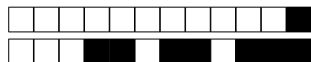
Avec les données `cout_prov.csv`, utilisez la fonction `aggregate` afin de trouver le coût moyen des sinistres par province. Ensuite, faites un graphique de type *bar plot* qui illustre le coût moyen des sinistres par province. Assurez-vous d'avoir les coûts de sinistre en ordre croissant.

Vous devriez alors obtenir le graphique suivant:

Veuillez noter que la couleur des bars n'est pas importante, mais si vous voulez obtenir la même couleur c'est le numéro `#e878b0`

N'oubliez pas que vous pouvez utiliser l'aide mémoire imprimé sur l'utilisation de `ggplot2`



**Question 7 (15 pts)**

Selon la nouvelle réforme de la sécurité routière au Québec, les détenteurs d'un permis probatoire pour les véhicules de promenade âgés de 19 ans et moins pourront transporter un nombre limité de passagers. Dans cet exercice, nous allons étudier l'impact de cette nouvelle réforme au sein de la compagnie d'assurance ACT3035. Votre superviseur vous demande alors de produire un graphique qui illustre les coûts de sinistre moyens pour les assurés âgés de 19 ans et moins, ainsi que pour tous les autres assurés. Afin d'y arriver, répondez à toutes les sous questions en

utilisant les deux ensembles de données `donnees_demo.csv` et `police_assurance.csv`.

a) (1 pts)

D'abord, quel est le nombre de personnes touchées par cette nouvelle réforme dans votre base de données?

b) (2 pts)

Dans votre *data frame* (df) 'police_assurance', créez une nouvelle variable appelée 'somme_couts' qui somme les coûts de sinistres (cout1, cout2, ..., cout7)

c) (2 pts)

En plus de la nouvelle variable créée en b), ajoutez une nouvelle variable 'sinistre_ann' qui extrait (ou transforme) l'année de la variable 'debut_pol' sous le format "%Y". Par exemple 'sinistre_ann' de la première ligne serait de 1999

d) (3 pts)

Pour cette question, vous pouvez utiliser le package `dplyr`

À partir de la base de données 'police_assurance' à laquelle vous avez ajouté les infos de la question b) et c), faites une jointure avec la base de données 'donnees_demo' afin d'avoir un tout nouveau df avec la variable `agee` provenant de la bd 'donnees_demo'.

e) (3 pts)

Dans le data frame de la question précédente, créez une nouvelle variable 'cat_age' qui divise les assurés en deux groupes selon leur âge;

- 19_et_moins
- Plus_que_19

Table 1: Data Frame final

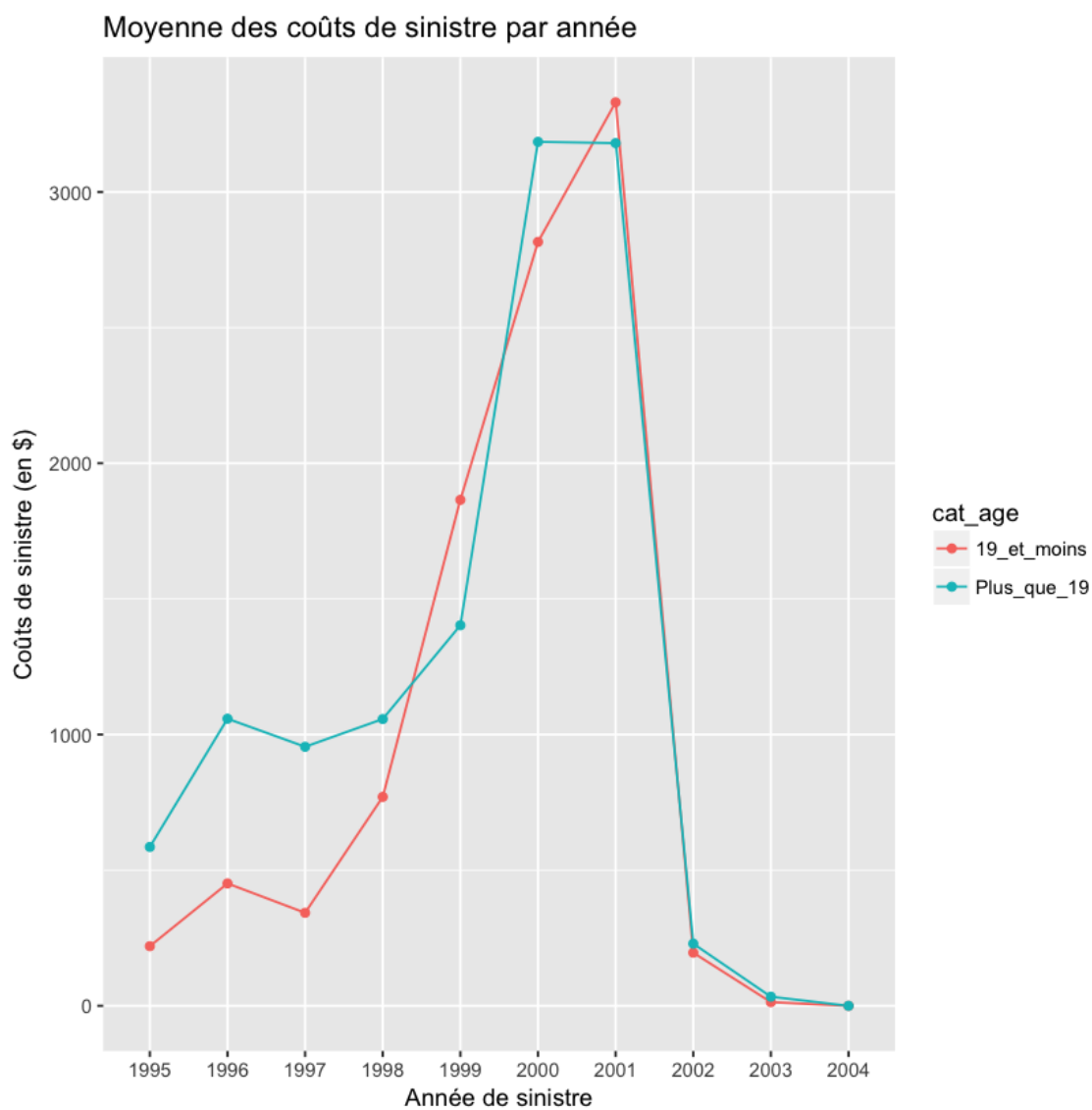
numeropol	debut_pol	fin_pol	nbsin	somme_couts	sinistre_ann	agee	cat_age
1	1999-11-10	2000-10-16	0	0.0000	1999	72	Plus_que_19
1	2000-10-17	2000-11-09	0	0.0000	2000	72	Plus_que_19
1	2000-11-10	2001-11-09	1	243.8571	2000	72	Plus_que_19
5	1996-01-03	1996-03-27	0	0.0000	1996	31	Plus_que_19
5	1996-03-28	1997-01-02	0	0.0000	1996	31	Plus_que_19
5	1997-01-03	1998-01-02	0	0.0000	1997	31	Plus_que_19



f) (4 pts)

Faites un graphique de deux lignes. La première représente la moyenne des coûts de sinistres des assurés de 19 ans et moins durant toutes les années d'accident. La deuxième ligne représente la moyenne des autres assurés âgés de plus de 19 ans.

Vous devriez obtenir un graphique tel qu'illustré à la figure ci-dessous;



**Question 8 (9 pts)**

Vous êtes nouvellement engagé à travailler au sein d'une petite compagnie d'assurance vie. On vous fait une petite formation lors de vos premiers jours et on vous présente les tâches de vos collègues. On vous dit que si vous avez besoin d'informations concernant les tâches administratives, vous pouvez vous adresser aux deux analystes en administration. Un des deux analystes vous explique qu'ils sont très occupés ces derniers jours, car ils doivent envoyer des lettres de proposition d'achat d'un nouveau produit d'assurances maladie graves aux clients. il vous explique ce qu'ils doivent faire pour y parvenir.

1. Ouvrir le *template* de la lettre
2. Remplacer le mot "LaDate" par la date du jour de rédaction de la lettre (change dans le temps)
3. Remplacer le mot "VotreNom" par "Martin Matte"
4. Lire le nom du client dans la base de données `clients` le nom du client, et remplace le mot "LeClient" dans le *template* par:
 - **Chère** Nom Client, pour les femmes, par exemple <Chère Anaïs-Adrienne Launay,>
 - **Cher** Nom Client, pour les hommes, par exemple <Cher Benjamin Delmas de la Gros,>
5. Enfin, ils enregistrent chaque lettre dans un dossier appelé "LettreAssuranceGraves" dans le répertoire Desktop de leur bureau. Chaque lettre porte le nom du client sans espace sous le format `.txt` comme suit: <Anaïs-AdrienneLaunay.txt>

Créer un script R que lorsque vous l'exécutez, les tâches de votre collègue *cute* sont exécuté en une fraction de seconde.

Faites votre propre code ou vous pouvez vous inspirer du code suivant *

```
clients<-read.csv("https://s3.amazonaws.com/www.nour.me/ACT3035/data/clients.csv") # pour lire votre base de donnees qui contient les noms et sexe des clients
tache_1 de votre collègue
...
...
for (i in ...){
  lettre_1<-readLines("https://s3.amazonaws.com/www.nour.me/ACT3035/data/lettre_2.txt") # pour lire le contenu du template de la lettre
  ...
  ...
  ...
  writeLines(lettre_2, con=gsub("_", "", nomFichier)) #pour sauvegarder votre fichier sous le nom nomFichier que vous avez défini
}
```

*à la dernière ligne de ce code, on remplace l'espace " " par "" aucun espace